

J.E.P.S.

1635

RECEIVED

AUG 28 2001 OIPE

TECH CENTER 1600/2900

#7

BP  
8-30-01

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/534,995

DATE: 07/24/2001

TIME: 10:56:15

Input Set : A:\021314310.txt

Output Set: N:\CRF3\07242001\I534995.raw

ENTERED

OK C-->

```
3 <110> APPLICANT: NISHIMURA, SATORU
4 KOIKE, AYUMI
6 <120> TITLE OF INVENTION: CHOLINE MONOOXYGENASE GENE
8 <130> FILE REFERENCE: 0213-1431-0
10 <140> CURRENT APPLICATION NUMBER: 09/534,995
11 <141> CURRENT FILING DATE: 2001-03-27
13 <150> PRIOR APPLICATION NUMBER: JP 273275/1999
14 <151> PRIOR FILING DATE: 1999-09-27
16 <160> NUMBER OF SEQ ID NOS: 17
18 <170> SOFTWARE: PatentIn version 3.1
20 <210> SEQ ID NO: 1
21 <211> LENGTH: 1828
22 <212> TYPE: DNA
23 <213> ORGANISM: Chenopodium album
25 <220> FEATURE:
26 <221> NAME/KEY: CDS
27 <222> LOCATION: (129)..(1427)
28 <223> OTHER INFORMATION:
31 <400> SEQUENCE: 1
32 gaactataca agctaagtta agcttaagct atattgttga tcattcttca tactacttcc 60
34 ttttaaaaaa aaattataac aacaaaagga agtgtgaatt ttttccttga tcatcatata 120
36 acatcaat atg gca gca agt gca aca aca atg ttg ctg aaa tac cca aca 170
37 Met Ala Ala Ser Ala Thr Thr Met Leu Leu Lys Tyr Pro Thr
38 1 5 10
40 act gta tgt ggt ata cca aat tca tca tca aac aat gat act tca aat 218
41 Thr Val Cys Gly Ile Pro Asn Ser Ser Ser Asn Asn Asp Thr Ser Asn
42 15 20 25 30
44 aat ata gtc cca att cca caa act agt act aat aat ccg gta ctt aag 266
45 Asn Ile Val Pro Ile Pro Gln Thr Ser Thr Asn Asn Pro Val Leu Lys
46 35 40 45
48 ttt cgt acc cct aat aaa acc att aac gcc gtc gct gcc ccg gct ttt 314
49 Phe Arg Thr Pro Asn Lys Thr Ile Asn Ala Val Ala Ala Pro Ala Phe
50 50 55 60
52 cct tct tta aac acc acc act act ccg ccg tcg att caa tca ctt gtc 362
53 Pro Ser Leu Asn Thr Thr Thr Thr Pro Pro Ser Ile Gln Ser Leu Val
54 65 70 75
56 cag gaa ttc gat ccg aag att ccg gct aag gat gct ctt acg cct cct 410
57 Gln Glu Phe Asp Pro Lys Ile Pro Ala Lys Asp Ala Leu Thr Pro Pro
58 80 85 90
60 agc tct tgg tat act gac gct gct ttc tat gct cat gaa ctt gac cgt 458
61 Ser Ser Trp Tyr Thr Asp Ala Ala Phe Tyr Ala His Glu Leu Asp Arg
62 95 100 105 110
64 atc ttt tat aag gga tgg caa gtc cca ggg tac agt gat caa att aag 506
65 Ile Phe Tyr Lys Gly Trp Gln Val Pro Gly Tyr Ser Asp Gln Ile Lys
66 115 120 125
68 gag cct aac caa tat ttc acc gga acg tta gga aat gtt gaa tat ttg 554
69 Glu Pro Asn Gln Tyr Phe Thr Gly Thr Leu Gly Asn Val Glu Tyr Leu
```

DATE: 07/24/2001

TIME: 10:56:15

Input Set : A:\021314310.txt

Output Set: N:\CRF3\07242001\I534995.raw

70				130				135				140					
72	gtg	tgt	cga	gat	ggg	gaa	gga	aaa	gtt	cat	gca	ttt	cac	aac	gtt	tgc	602
73	Val	Cys	Arg	Asp	Gly	Glu	Gly	Lys	Val	His	Ala	Phe	His	Asn	Val	Cys	
74			145					150					155				
76	acc	cat	cgt	gct	tcg	att	ctt	gct	tgt	gga	agt	gga	aaa	aaa	tcg	tgt	650
77	Thr	His	Arg	Ala	Ser	Ile	Leu	Ala	Cys	Gly	Ser	Gly	Lys	Lys	Ser	Cys	
78		160					165					170					
80	ttt	gtg	tgc	cct	tac	cat	gga	tgg	gta	ttt	ggc	atg	aat	gga	tcg	ctt	698
81	Phe	Val	Cys	Pro	Tyr	His	Gly	Trp	Val	Phe	Gly	Met	Asn	Gly	Ser	Leu	
82	175					180					185				190		
84	aca	aaa	gct	tcc	aaa	gca	acc	gaa	gaa	cag	tca	ctt	gat	ccc	gat	gaa	746
85	Thr	Lys	Ala	Ser	Lys	Ala	Thr	Glu	Glu	Gln	Ser	Leu	Asp	Pro	Asp	Glu	
86				195						200				205			
88	ctt	ggg	ctt	gta	ccc	ctg	aaa	gtt	gca	gta	tgg	ggc	cca	ttt	ata	ctc	794
89	Leu	Gly	Leu	Val	Pro	Leu	Lys	Val	Ala	Val	Trp	Gly	Pro	Phe	Ile	Leu	
90			210					215					220				
92	ata	agt	ttg	gac	aga	tca	agc	ctt	gaa	gta	ggg	gat	gtt	gga	tct	gaa	842
93	Ile	Ser	Leu	Asp	Arg	Ser	Ser	Leu	Glu	Val	Gly	Asp	Val	Gly	Ser	Glu	
94		225					230					235					
96	tgg	ctt	ggg	agt	tgt	gct	gaa	gat	gtt	aag	gcc	cat	gct	ttt	gac	cct	890
97	Trp	Leu	Gly	Ser	Cys	Ala	Glu	Asp	Val	Lys	Ala	His	Ala	Phe	Asp	Pro	
98		240				245					250						
100	aat	tta	cag	ttc	atc	aat	agg	agt	gaa	ttt	cca	atg	gaa	tct	aat	tgg	938
101	Asn	Leu	Gln	Phe	Ile	Asn	Arg	Ser	Glu	Phe	Pro	Met	Glu	Ser	Asn	Trp	
102	255				260						265				270		
104	aag	att	ttc	agt	gac	aac	tat	ttg	gat	agc	tcg	tac	cat	gtt	cct	tat	986
105	Lys	Ile	Phe	Ser	Asp	Asn	Tyr	Leu	Asp	Ser	Ser	Tyr	His	Val	Pro	Tyr	
106				275						280				285			
108	gca	cac	aag	tac	tat	gct	act	gaa	ctc	gac	ttt	gat	act	tac	caa	act	1034
109	Ala	His	Lys	Tyr	Tyr	Ala	Thr	Glu	Leu	Asp	Phe	Asp	Thr	Tyr	Gln	Thr	
110			290					295					300				
112	gat	atg	atc	gga	aac	gtc	acg	att	caa	aga	gtg	gca	ggg	agt	tca	aac	1082
113	Asp	Met	Ile	Gly	Asn	Val	Thr	Ile	Gln	Arg	Val	Ala	Gly	Ser	Ser	Asn	
114		305				310						315					
116	aat	ggg	ttt	aat	aga	ctt	gga	tct	caa	gca	ttc	tat	gct	ttt	gca	tac	1130
117	Asn	Gly	Phe	Asn	Arg	Leu	Gly	Ser	Gln	Ala	Phe	Tyr	Ala	Phe	Ala	Tyr	
118		320				325					330						

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/534,995

DATE: 07/24/2001

TIME: 10:56:16

Input Set : A:\021314310.txt

Output Set: N:\CRF3\07242001\I534995.raw

```

136 gtc caa aaa ggg ttg gaa aca cca gca tat cgt agt gga aga tat gtg      1370
137 Val Gln Lys Gly Leu Glu Thr Pro Ala Tyr Arg Ser Gly Arg Tyr Val
138      400      405      410
140 atg cca att gag aaa gga atc cat cat ttc cac tgc tgg ttg cac caa      1418
141 Met Pro Ile Glu Lys Gly Ile His His Phe His Cys Trp Leu His Gln
142 415      420      425      430
144 gta ttg aag tgattgcagc agatcatcag atgttcgttt cttcttgat      1467
145 Val Leu Lys
148 tggaattgga tattatgatt aataagtaaa attataatgt cataatgtag ttgagattgt      1527
150 tgctagagtt gagcgtatgc tcctcatgca ctacttagtt atcaagtgtg tatgtctttg      1587
152 gtcatgggca aaatgtatgt ttcttgctag aatttatata ttatgggtgt aatgtccaat      1647
154 ataaataaaa accatagcac ccctttaatt ccctacttag gtttatatcc catttatitt      1707
156 cgggggatct atgagataga ttgtctatga acattatitt tcgactcgtg tatggtatcc      1767
158 atcccttggt taggggtgaag taaacattga gtgtatgaag ttttcattga gtttctgctt      1827
160 t      1828
163 <210> SEQ ID NO: 2
164 <211> LENGTH: 433
165 <212> TYPE: PRT
166 <213> ORGANISM: Chenopodium album
168 <400> SEQUENCE: 2
170 Met Ala Ala Ser Ala Thr Thr Met Leu Leu Lys Tyr Pro Thr Thr Val
171 1      5      10      15
174 Cys Gly Ile Pro Asn Ser Ser Ser Asn Asn Asp Thr Ser Asn Asn Ile
175      20      25      30
178 Val Pro Ile Pro Gln Thr Ser Thr Asn Asn Pro Val Leu Lys Phe Arg
179      35      40      45
182 Thr Pro Asn Lys Thr Ile Asn Ala Val Ala Ala Pro Ala Phe Pro Ser
183      50      55      60
186 Leu Asn Thr Thr Thr Thr Pro Pro Ser Ile Gln Ser Leu Val Gln Glu
187 65      70      75      80
190 Phe Asp Pro Lys Ile Pro Ala Lys Asp Ala Leu Thr Pro Pro Ser Ser
191      85      90      95
194 Trp Tyr Thr Asp Ala Ala Phe Tyr Ala His Glu Leu Asp Arg Ile Phe
195      100      105      110
198 Tyr Lys Gly Trp Gln Val Pro Gly Tyr Ser Asp Gln Ile Lys Glu Pro
199      115      120      125
202 Asn Gln Tyr Phe Thr Gly Thr Leu Gly Asn Val Glu Tyr Leu Val Cys
203      130      135      140
206 Arg Asp Gly Glu Gly Lys Val His Ala Phe His Asn Val Cys Thr His
207 145      150      155      160
210 Arg Ala Ser Ile Leu Ala Cys Gly Ser Gly Lys Lys Ser Cys Phe Val
211      165      170      175
214 Cys Pro Tyr His Gly Trp Val Phe Gly Met Asn Gly Ser Leu Thr Lys
215      180      185      190
218 Ala Ser Lys Ala Thr Glu Glu Gln Ser Leu Asp Pro Asp Glu Leu Gly
219      195      200      205
222 Leu Val Pro Leu Lys Val Ala Val Trp Gly Pro Phe Ile Leu Ile Ser
223      210      215      220
226 Leu Asp Arg Ser Ser Leu Glu Val Gly Asp Val Gly Ser Glu Trp Leu

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/534,995

DATE: 07/24/2001

TIME: 10:56:16

Input Set : A:\021314310.txt

Output Set: N:\CRF3\07242001\I534995.raw

```

227 225          230          235          240
230 Gly Ser Cys Ala Glu Asp Val Lys Ala His Ala Phe Asp Pro Asn Leu
231          245          250          255
234 Gln Phe Ile Asn Arg Ser Glu Phe Pro Met Glu Ser Asn Trp Lys Ile
235          260          265          270
238 Phe Ser Asp Asn Tyr Leu Asp Ser Ser Tyr His Val Pro Tyr Ala His
239          275          280          285
242 Lys Tyr Tyr Ala Thr Glu Leu Asp Phe Asp Thr Tyr Gln Thr Asp Met
243          290          295          300
246 Ile Gly Asn Val Thr Ile Gln Arg Val Ala Gly Ser Ser Asn Asn Gly
247 305          310          315          320
250 Phe Asn Arg Leu Gly Ser Gln Ala Phe Tyr Ala Phe Ala Tyr Pro Asn
251          325          330          335
254 Phe Ala Val Glu Arg Tyr Gly Pro Trp Met Thr Thr Met His Ile Leu
255          340          345          350
258 Pro Leu Gly Pro Arg Lys Cys Lys Leu Val Val Asp Tyr Tyr Ile Glu
259          355          360          365
262 Lys Ser Lys Leu Asp Asp Lys Asp Tyr Ile Glu Lys Gly Ile Ala Ile
263          370          375          380
266 Asn Asp Asn Val Gln Lys Glu Asp Val Val Leu Cys Glu Ser Val Gln
267 385          390          395          400
270 Lys Gly Leu Glu Thr Pro Ala Tyr Arg Ser Gly Arg Tyr Val Met Pro
271          405          410          415
274 Ile Glu Lys Gly Ile His His Phe His Cys Trp Leu His Gln Val Leu
275          420          425          430
278 Lys
282 <210> SEQ ID NO: 3
283 <211> LENGTH: 1651
284 <212> TYPE: DNA
285 <213> ORGANISM: Chenopodium album
287 <220> FEATURE:
288 <221> NAME/KEY: CDS
289 <222> LOCATION: (119)..(1423)
290 <223> OTHER INFORMATION:
293 <400> SEQUENCE: 3
294 cttgaattat acaagctaag tatatatatt gttgatcatc tttcatacca cctttaaaaa 60
296 aaattataac aacaaaagga agtggttagt tattgcttga tcatcatata atatcaac 118
298 atg tca gca agt gca aca aca atg ttg ctg aaa tac cca aca act gta 166
299 Met Ser Ala Ser Ala Thr Thr Met Leu Leu Lys Tyr Pro Thr Thr Val
300 1          5          10          15
302 tgt ggt ata cca aat tca tca tca aac aat gat act tca aat aac atc 214
303 Cys Gly Ile Pro Asn Ser Ser Ser Asn Asn Asp Thr Ser Asn Asn Ile
304          20          25          30
306 gtc cca att cca caa act agt act aat aat ccg gta ctt aag ttt cgt 262
307 Val Pro Ile Pro Gln Thr Ser Thr Asn Asn Pro Val Leu Lys Phe Arg
308          35          40          45
310 acc cct aat aaa acc att aac gcc gtc gct gcc ccg gct ttt cct tct 310
311 Thr Pro Asn Lys Thr Ile Asn Ala Val Ala Ala Pro Ala Phe Pro Ser
312          50          55          60

```

## RAW SEQUENCE LISTING

DATE: 07/24/2001

PATENT APPLICATION: US/09/534,995

TIME: 10:56:16

Input Set : A:\021314310.txt

Output Set: N:\CRF3\07242001\I534995.raw

314	tta	agt	acc	acc	act	act	ccg	ccg	tcg	att	caa	tca	ctt	gtc	cag	gaa	358
315	Leu	Ser	Thr	Thr	Thr	Thr	Pro	Pro	Ser	Ile	Gln	Ser	Leu	Val	Gln	Glu	
316	65					70					75					80	
318	ttc	gat	ccg	agg	att	ctg	gcc	gag	gat	gct	ctc	acg	cct	cct	agc	tct	406
319	Phe	Asp	Pro	Arg	Ile	Leu	Ala	Glu	Asp	Ala	Leu	Thr	Pro	Pro	Ser	Ser	
320					85					90					95		
322	tgg	tat	act	gaa	cct	gcc	ttc	tat	gct	cat	gaa	ctt	gac	cgt	atc	ttt	454
323	Trp	Tyr	Thr	Glu	Pro	Ala	Phe	Tyr	Ala	His	Glu	Leu	Asp	Arg	Ile	Phe	
324				100					105					110			
326	tac	aaa	gga	tgg	caa	gtc	gca	ggg	tac	agc	gat	caa	att	aag	gag	cct	502
327	Tyr	Lys	Gly	Trp	Gln	Val	Ala	Gly	Tyr	Ser	Asp	Gln	Ile	Lys	Glu	Pro	
328			115					120					125				
330	aac	caa	tat	ttc	acc	gga	acg	tta	gga	aat	gtt	gaa	tat	ttg	gtg	tgt	550
331	Asn	Gln	Tyr	Phe	Thr	Gly	Thr	Leu	Gly	Asn	Val	Glu	Tyr	Leu	Val	Cys	
332		130					135				140						
334	cga	gat	ggt	gaa	gga	aaa	gtt	cat	gca	ttt	cac	aat	gtt	tgc	act	cat	598
335	Arg	Asp	Gly	Glu	Gly	Lys	Val	His	Ala	Phe	His	Asn	Val	Cys	Thr	His	
336	145					150				155						160	
338	cgt	gct	tcg	att	ctt	gct	tgt	gga	agt	ggc	aaa	aaa	tcg	tgt	ttc	gta	646
339	Arg	Ala	Ser	Ile	Leu	Ala	Cys	Gly	Ser	Gly	Lys	Lys	Ser	Cys	Phe	Val	
340					165					170					175		
342	tgc	cct	tac	cat	ggt	tgg	gta	ttt	ggc	atg	aat	gga	tca	ctt	acg	aaa	694
343	Cys	Pro	Tyr	His	Gly	Trp	Val	Phe	Gly	Met	Asn	Gly	Ser	Leu	Thr	Lys	
344				180					185					190			
346	gct	tcc	aaa	gca	acc	gaa	gaa	cag	tcc	ctt	gat	ccc	gat	gaa	ctt	ggg	742
347	Ala	Ser	Lys	Ala	Thr	Glu	Glu	Gln	Ser	Leu	Asp	Pro	Asp	Glu	Leu	Gly	
348			195					200					205				
350	ctt	gta	ccc	ctg	aaa	gtt	gca	gta	tgg	ggc	cca	ttt	ata	ctc	atc	agt	790
351	Leu	Val	Pro	Leu	Lys	Val	Ala	Val	Trp	Gly	Pro	Phe	Ile	Leu	Ile	Ser	
352		210					215				220						
354	ttg	gac	aga	tca	agc	ctt	gaa	gta	ggc	gat	gtt	gga	tct	gaa	tgg	ctt	838
355	Leu	Asp	Arg	Ser	Ser	Leu	Glu	Val	Gly	Asp	Val	Gly	Ser	Glu	Trp	Leu	
356	225					230				235					240		
358	ggt	agt	tgt	gct	gaa	gat	gtt	aag	gcc	cat	gct	ttt	gac	cct	aat	ttg	886
359	Gly	Ser	Cys	Ala	Glu	Asp	Val	Lys	Ala	His	Ala	Phe	Asp	Pro	Asn	Leu	
360					245				250						255		
362	cag	ttc	atc	aat	agg	agt	gaa	ttt	cca	atg	gaa	tct	aat	tgg	aag	att	934
363	Gln	Phe	Ile	Asn	Arg	Ser	Glu	Phe	Pro	Met	Glu	Ser	Asn	Trp	Lys	Ile	
364				260					265					270			
366	ttc	agt	gac	aac	tac	ttg	gat	agc	tcg	tac	cat	gtt	cct	tat	gca	cac	982
367	Phe	Ser	Asp	Asn	Tyr	Leu	Asp	Ser	Ser	Tyr	His	Val	Pro	Tyr	Ala	His	
368			275					280					285				
370	aag	tac	tat	gca	act	gaa	ctc	gac	ttt	gat	act	tat	caa	acc	gat	atg	1030
371	Lys	Tyr	Tyr	Ala	Thr	Glu	Leu	Asp	Phe	Asp	Thr	Tyr	Gln	Thr	Asp	Met	
372		290					295				300						
374	att	gga	aat	gtc	acg	att	caa	aga	gtg	gcg	ggg	agt	tca	aac	aag	cca	1078
375	Ile	Gly	Asn	Val	Thr	Ile	Gln	Arg	Val	Ala	Gly	Ser	Ser	Asn	Lys	Pro	
376	305					310				315					320		
378	gat	ggt	ttt	gat	aga	ctt	gga	tct	caa	gca	ttc	tat	gct	ttt	gca	tac	1126

**Please Note:**

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

**VERIFICATION SUMMARY**

PATENT APPLICATION: US/09/534,995

DATE: 07/24/2001

TIME: 10:56:17

Input Set : A:\021314310.txt

Output Set: N:\CRF3\07242001\I534995.raw

L:11 M:271 C: Current Filing Date differs, Replaced Current Filing Date

L:823 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7

L:853 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8